

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Faculty Papers and Publications in Animal
Science

Animal Science Department

2-15-2021

Generation of a Biobank From Two Adult Thoroughbred Stallions for the Functional Annotation of Animal Genomes Initiative

Callum G. Donnelly

Rebecca R. Bellone

Erin N. Hales

Annee Nguyen

Scott A. Katzman

See next page for additional authors

Follow this and additional works at: <https://digitalcommons.unl.edu/animalscifacpub>



Part of the [Genetics and Genomics Commons](#), and the [Meat Science Commons](#)

This Article is brought to you for free and open access by the Animal Science Department at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Papers and Publications in Animal Science by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Authors

Callum G. Donnelly, Rebecca R. Bellone, Erin N. Hales, Annee Nguyen, Scott A. Katzman, Ghislaine A. Dujovne, Kelly E. Knickelbein, Felipe Avila, Ted S. Kalbfleisch, Elena Giulotto, Nicole B. Kingsley, Jocelyn Tanaka, Elizabeth Esdaile, Sichong Peng, Anna Dahlgren, Anna Fuller, Michael J. Mienaltowski, Terje Raudsepp, Verena K. Affolter, Jessica L. Petersen, and Carrie J. Finno



Open Access CC-BY

Generation of a Biobank From Two Adult Thoroughbred Stallions for the Functional Annotation of Animal Genomes Initiative

Callum G. Donnelly¹, Rebecca R. Bellone^{1,2}, Erin N. Hales³, Annee Nguyen¹, Scott A. Katzman⁴, Ghislaine A. Dujovne¹, Kelly E. Knickelbein⁴, Felipe Avila², Ted S. Kalbfleisch⁵, Elena Giulotto⁶, Nicole B. Kingsley², Jocelyn Tanaka², Elizabeth Esdaile², Sichong Peng¹, Anna Dahlgren¹, Anna Fuller⁷, Michael J. Mienaltowski⁸, Terje Raudsepp⁹, Verena K. Affolter¹⁰, Jessica L. Petersen⁷ and Carrie J. Finno^{1*}

OPEN ACCESS

Edited by:

Hans Cheng,
Agricultural Research Service (USDA),
United States

Reviewed by:

Tosso Leeb,
University of Bern, Switzerland
Patrick Venta,
Michigan State University,
United States

*Correspondence:

Carrie J. Finno
cjfinno@ucdavis.edu

Specialty section:

This article was submitted to
Livestock Genomics,
a section of the journal
Frontiers in Genetics

Received: 06 January 2021

Accepted: 15 February 2021

Published: 08 March 2021

Citation:

Donnelly CG, Bellone RR, Hales EN, Nguyen A, Katzman SA, Dujovne GA, Knickelbein KE, Avila F, Kalbfleisch TS, Giulotto E, Kingsley NB, Tanaka J, Esdaile E, Peng S, Dahlgren A, Fuller A, Mienaltowski MJ, Raudsepp T, Affolter VK, Petersen JL and Finno CJ (2021) Generation of a Biobank From Two Adult Thoroughbred Stallions for the Functional Annotation of Animal Genomes Initiative. *Front. Genet.* 12:650305. doi: 10.3389/fgene.2021.650305

¹ Department of Population Health and Reproduction, School of Veterinary Medicine, University of California, Davis, Davis, CA, United States, ² Veterinary Genetics Laboratory, School of Veterinary Medicine, University of California, Davis, Davis, CA, United States, ³ Morris Animal Foundation, Denver, CO, United States, ⁴ Department of Surgical and Radiological Sciences, School of Veterinary Medicine, University of California, Davis, Davis, CA, United States, ⁵ Gluck Equine Research Center, University of Kentucky, Lexington, KY, United States, ⁶ Department of Biology and Biotechnology, University of Pavia, Pavia, Italy, ⁷ Department of Animal Science, University of Nebraska-Lincoln, Lincoln, NE, United States, ⁸ Department of Animal Science, College of Agricultural and Environmental Sciences, University of California, Davis, Davis, CA, United States, ⁹ Department of Veterinary Integrative Biosciences, Texas A&M University, College Station, TX, United States, ¹⁰ Department of Pathology, Microbiology, and Immunology, School of Veterinary Medicine, University of California, Davis, Davis, CA, United States

Following the successful creation of a biobank from two adult Thoroughbred mares, this study aimed to recapitulate sample collection in two adult Thoroughbred stallions as part of the Functional Annotation of the Animal Genome (FAANG) initiative. Both stallions underwent thorough physical, lameness, neurologic, and ophthalmic (including electroretinography) examinations prior to humane euthanasia. Epididymal sperm was recovered from both stallions immediately postmortem and cryopreserved. Aseptically collected full thickness skin biopsies were used to isolate, culture and cryopreserve dermal fibroblasts. Serum, plasma, cerebrospinal fluid, urine, and gastrointestinal content from various locations were collected and cryopreserved. Under guidance of a board-certified veterinary anatomic pathologist, 102 representative tissue samples were collected from both horses. Whole tissue samples were flash-frozen and prioritized tissues had nuclei isolated and cryopreserved. Spatially contemporaneous samples of each tissue were submitted for histologic examination. Antemortem and gross pathologic examination revealed mild abnormalities in both stallions. One stallion (ECA_UCD_AH3) had unilateral thoracic limb lameness and bilateral chorioretinal scars. The second stallion (ECA_UCD_AH4) had subtle symmetrical pelvic limb ataxia, symmetrical prostatomegaly, and moderate gastrointestinal nematodiasis. DNA from each was whole-genome sequenced and genotyped using the GGP Equine 70K SNP array. The genomic resources and banked biological samples from these animals

augments the existing resource available to the equine genomics community. Importantly we may now improve the resolution of tissue-specific gene regulation as affected by sex, as well as add sex-specific tissues and gametes.

Keywords: FAANG, horse, male, stallion, biobank

INTRODUCTION

Sex bias in animal model-based biomedical research has been a focus area for the National Institutes for Health (NIH) (Will et al., 2017). While the influence of sex on experimental outcomes is undeniable, the adoption of sex as a consideration for experimental design has been slow (Zucker and Beery, 2010). Genomic regulation is particularly labile to the effect of sex (Lee, 2018). With the intent of cataloging the functional and regulatory elements of economically significant animal species, there is need for the Functional Annotation of Animal Genome (FAANG) project to diligently account for sex differences. This is of importance to the equine industry in which sex of animal is an important component of management and use, such as predominance of females in polo competition and males in show jumping (Fenner et al., 2019). The increased resolution of tissue specific genomic, transcriptional, and epigenomic data will be of significant benefit to the progress of equine genomics. In particular, the ability to more precisely characterize regulation of genomic regions associated with traits pertinent to equine health and performance are made increasingly more possible (Kingsley et al., 2019). The objective of the current study was to augment the existing equine biobank with male samples. We report here the addition of two male Thoroughbred horses to the biobank of two female Thoroughbred horses initially established in 2016.

MATERIALS AND METHODS

Animals

Two Thoroughbred stallions (aged 3 and 4 years; ECA_UCD_AH3 and ECA_UCD_AH4, respectively) were donated for this project. ECA_UCD_AH3 had been in racing training and sustained a career ending musculoskeletal injury prior to donation. ECA_UCD_AH4 had not engaged in racing training, and is a son of the reference genome donor Twilight. Approval for all protocols was granted by the UC Davis Institutional Animal Care and Use Committee (Protocol #21033).

Clinical Assessments

Complete physical examination (including complete blood count and serum biochemistry) was performed and interpreted by two board-certified internists (CJF and CGD) for both stallions. Complete neurologic and lameness examinations were performed by relevant recognized experts (CJF and SAK, respectively) for both stallions within 24 h prior to euthanasia. Complete ophthalmic examination of the anterior and posterior segment (performed by KEK) as well as photopic and scotopic electroretinography (ERG) was performed on both stallions within 48 h prior to euthanasia.

Clinical Sample Collection

Serum and plasma samples were collected immediately prior to euthanasia *via* an indwelling intravenous cannula. Heparin and EDTA whole blood was centrifuged at $2,000\times g$ for 10 min at 4°C . Plasma was harvested, flash frozen in liquid nitrogen and stored at -80°C . Whole blood in plain tubes was allowed to clot for 30 min at room temperature, with serum collected in the same fashion as plasma.

Ejaculated sperm and seminal plasma were collected from ECA_UCD_AH3 at the time of ophthalmic examination. Ejaculated sperm was unable to be collected from ECA_UCD_AH4. Seminal plasma was retrieved by uncushioned centrifugation at $2,000\times g$ for 15 min. The supernatant was separated and centrifuged again before being passed through a $20\mu\text{m}$ filter. The supernatant and sperm pellet were flash frozen in liquid nitrogen and stored at -80°C .

Immediately following euthanasia, the left testicle, epididymis, and distal ductus deferens were removed from each horse. The epididymes were removed and incised sharply transversely at the cauda epididymis. Commercial semen extender was flushed retrograde *via* the remaining ductus deferens. Recovered spermatozoa were frozen in 0.5 mL polyethylene straws by controlled rate freezing protocol and stored in liquid nitrogen, at 323 and 455.5 million sperm/mL for ECA_UCD_AH3 and ECA_UCD_AH4, respectively.

Synovial fluid was collected aseptically (by centesis) from the left middle carpal, radio-carpal joint and the left medial compartment of the femoro-tibial joint from ECA_UCD_AH3 and the right middle carpal, radio-carpal joint and left medial compartment of the femoro-tibial joint from ECA_UCD_AH4 immediately following euthanasia. A veterinary clinical pathologist cytologically evaluated an aliquot of each sample within 2 h of collection. The remaining sample was centrifuged at $2,000\times g$ for 10 min at 4°C , and the supernatant was flash frozen in liquid nitrogen and stored at -80°C .

Cerebrospinal fluid was collected aseptically *via* atlanto-occipital centesis. A veterinary clinical pathologist cytologically evaluated an aliquot of each sample within 2 h of collection. The remaining sample was then processed in the same manner as synovial fluid.

Peripheral Blood Mononuclear Cell Collection

Whole blood was collected in EDTA tubes 24 h prior to humane euthanasia. Using density gradient centrifugation, peripheral blood mononuclear cells (PBMCs) were harvested as previously described (Hida et al., 2002). Briefly, whole blood was overlaid on a double gradient of 1,077 and 1,119 histopaque media (Sigma-Aldrich). Then, samples were centrifuged at $700 g$ for

30 min with the enriched PBMC layer carefully removed, flash frozen in liquid nitrogen and stored at -80°C .

Tissue Specific Sampling

Full thickness skin biopsies were aseptically collected from the area overlying the left gluteal muscles for dermal fibroblast isolation and cultured as previously described (Raimondi et al., 2011). Briefly, biopsies were washed in ice cold PBS with the addition of penicillin and streptomycin. Fragments of the dermis ($2\text{--}3\text{ mm}^2$) were placed into a 24-well tissue-culture-treated plate and covered in complete media (Dulbecco's Minimum Essential Medium, 20% fetal bovine serum, $2\times$ non-essential amino acids, 2 mM L-glutamine, $2\times$ penicillin/streptomycin, $2\text{ }\mu\text{g/ml}$ amphotericin B, and $1\text{ }\mu\text{g/ml}$ fluconazole). At confluence, cells were trypsinized, counted, and seeded in a 12-well plate. Cells were passaged and frozen at passage three and four in a DMSO-based cryoprotectant media and stored in liquid nitrogen (Raimondi et al., 2011). DNA isolated from whole blood and from cultured fibroblast cells of each horse were genotyped and compared for 15 standard genetic markers (fourteen microsatellites and one sex link marker amelogenin) routinely tested for use in identify and parentage testing services as the UC Davis Veterinary Genetics Laboratory.

A total of 102 issue samples were taken from all body systems (Supplementary Table 1). Sample stations were arranged in teams, with a veterinarian overseeing each team to ensure appropriate tissue identification. Additionally, all tissues were examined by a veterinary anatomic pathologist (VKA) for gross abnormalities prior to collection (Supplementary Table 2). Tissue samples were collected as previously described (Burns et al., 2018). Briefly, samples from the most representative portion of each tissue were collected and preserved in 10% buffered formalin for histopathology. Samples for nuclei preparation and tissue banking were taken from sites immediately adjacent (proximal and distal) to the representative histopathology sections and flash frozen. Nuclei isolation was performed as similarly to that described for the mare biobank, with the exceptions of testis replacing ovary and C6 spinal cord instead of T1 (Burns et al., 2018).

Genotyping and Whole-Genome Sequencing

Genomic DNA was isolated from whole blood using a previously validated technique (Burns et al., 2018). Whole-genome sequencing using an Illumina NovaSeq (150 bp paired-end reads) platform was performed at Admera Health, LLC (South Plainfield, NJ), targeting 35X coverage for each horse. Library preparation was performed using the KAPA library quantification kit for Illumina (Roche Holding AG, Basel, Switzerland). SNP genotyping was performed using the GGP Equine 70K SNP bead chip array (Neogen GeneSeek, Lincoln, NE). Genotype data were merged with those collected from the prior sampling effort of Thoroughbred mares (Burns et al., 2018). Marker positions were translated from EquCab2 to EquCab3 using the NCBI Genome Remapping Service (<https://www.ncbi.nlm.nih.gov/genome/tools/remap>), removing markers that were not mapped to any of the 31 autosomes. Using SNP data,

runs of homozygosity (ROH), a measure of diversity, was quantified in DetectRuns (<https://github.com/bioinformatics-ptp/detectRUNS/tree/master/detectRUNS>) using the method of Marras et al. (2015). Analysis parameters required a minimum of 15 SNPs, maxOpp Run = 1, max MissRun = 1, maxGap = 10,000,000, minLengthBps = 100,000. Whole-genome sequence data of the stallions and mares were processed and variants called according to the pipeline outlined in Sieck et al. (2020), with the exception of mapping to the EquCab3 reference genome (Burns et al., 2018; Sieck et al., 2020). Bi-allelic autosomal SNPs were extracted from the resulting vcf for ROH analysis in the same manner as used for the SNP data.

Additionally, in an effort to provide extensive phenotyping and genotyping for future studies on pigmentation biology, both stallions were genotyped for coat color loci using the commercially available horse full coat color and white pattern panel (<https://vgl.ucdavis.edu/panel/full-coat-color-pattern-panel>).

Pedigree Analysis

Individual inbreeding coefficients (F) were calculated from five-generation pedigrees of both stallions and the previously sampled mares (Burns et al., 2018) using Pedigraph (Garbe and Da, 2008).

Karyotyping

Sodium heparin stabilized whole blood samples were collected 7 days prior to euthanasia for karyotype characterization. Karyotypes were generated for each horse using the previously validated Pokeweed-stimulated lymphocyte culture (Raudsepp et al., 2010). Thirty metaphase cells from ECA_UCD_AH3 and ECA_UCD_AH4 were captured for analysis. Genomic DNA was isolated from lymphocyte cultures for PCR detection of the sex determining region Y (SRY) and androgen receptor (AR) genes, as previously described (Raudsepp et al., 2004).

RESULTS

Clinical Assessments

ECA_UCD_AH3 had several mild abnormalities apparent on clinical examination, including a grade III/V left thoracic limb lameness [AAEP scale (Baxter, 2020)] and chorioretinal scarring of both eyes. Scotopic ERG data for both horses is provided in Supplementary Table 3. No other abnormalities were detected antemortem for ECA_UCD_AH3. ECA_UCD_AH4 did not have clinically detectable lameness, however, a mild grade I/V [Modified Mayhew scale (Furr and Reed, 2015)] symmetrical pelvic limb ataxia was detected with a presumptive neuroanatomical localization to the cervical spinal cord. No other abnormalities were detected for ECA_UCD_AH4 on clinical examination. No abnormalities were detected on complete blood count or serum biochemistry for either stallion.

Clinical Sample Collection

Serum, plasma (collected in EDTA and heparin tubes), buffy coat and urine were collected from both stallions. Seminal plasma and ejaculated spermatozoa were obtained for ECA_UCD_AH3 only. Epididymal recovered spermatozoa from both stallions

were successfully isolated and cryopreserved with adequate post thaw kinematic parameters (**Supplementary Table 4**). Synovial and cerebrospinal fluid samples were cytologically normal (**Supplementary Tables 5, 6**). PBMCs were successfully collected from both stallions.

Genotyping

Achieved coverage of the whole-genome sequence data was 38.7 and 37.5X for ECA_UCD_AH3 and ECA_UCD_AH4, respectively. Relative to the reference genome, between 5.4 (ECA_UCD_AH3) and 3.8 (ECA_UCD_AH4) million variants were observed. Data are publically available through the FAANG consortium. Combining the 70K genotype data of the stallions to that from the two mares and removing variants present on sex chromosome data and on unmapped contigs, 59,823 variants remained. The count of ROH per horse was similar (498–503) for all four horses, except ECA_UCD_AH4 (280). However, ECA_UCD_AH4 had a greater number of long (>8 Mb) ROH than all other horses (**Supplementary Table 7**). Considering SNP data from WGS, between 1,822 (ECA_UCD_AH1) and 2,867 (ECA_UCD_AH4) ROH were observed. Five-generation pedigree-based inbreeding estimates ranged from 0 (ECA_UCD_AH1) to 0.344 (ECA_UCD_AH4).

Genetic profiles from cultured fibroblast cells matched to that of the expected horse (15/15 markers matched for each horse, presented in **Supplementary Table 8**).

ECA_UCD_AH3 is genetically defined as a bay horse (E/e A/a) with one copy of dominant white 20 (N/W20), while ECA_UCD_AH4 was gray with a bay base coat (E/e A/A) before graying out and was homozygous for dominant white 20 (W20/W20) (all coat color genotypes are presented in **Supplementary Table 9**).

Karyotype

Both stallions had a normal 64 XY male karyotype with no apparent structural or numerical chromosomal abnormalities. Both stallions were positive for the SRY and AR genes.

Tissue Specific Sampling

Sample collection commenced 30 min following euthanasia and was concluded within 3 h for all tissue samples. Samples from 102 tissues were collected and flash frozen (**Supplementary Table 1**). Sixteen of these tissues also had nuclei isolated and cryopreserved (**Supplementary Table 1**).

Pathology

Gross examination of ECA_UCD_AH3 was without abnormality, except for a 15 × 2 × 1.3 cm firm area at the caudal margin of the left caudal lung lobe. This area was therefore avoided for sampling. Histologic examination of the area identified locally extensive fibrosis, Type II pneumocyte proliferation and alveolar histiocytosis, which is consistent with a previous pulmonary inflammatory insult. Mild lymphocytic inflammation was noted at several sections of the gastrointestinal tract, but were considered incidental. Gross examination of ECA_UCD_AH4 revealed marked gastrointestinal nematodiasis throughout the large and small

intestinal segments. Gastrointestinal segments had marked lymphocytic and eosinophilic infiltrates consistent with gastrointestinal parasitism. ECA_UCD_AH4 had marked prostatomegally grossly, but was without abnormality on histopathology. There were no gross or histologic findings of the central nervous system of ECA_UCD_AH4 that could explain the mild sensory ataxia. There were no other significant histologic abnormalities. Summary of all pathologic findings are summarized in **Supplementary Table 2**.

DISCUSSION

We report the successful completion of the male equine biobank for the FAANG consortium. This project has added 102 tissues from two stallions in addition to seven body fluids, one cell line and spermatozoa to the biobank for use by the equine research community. This study recapitulated the guidelines of the original equine tissue collection, demonstrating that this method is repeatable and appropriate for this type of endeavor.

Stringent phenotype information both ante- and postmortem are a critical feature of this biobank and lends strength to the interpretation of results of subsequent studies using these tissues. Similarly, it also establishes the boundaries for the interpretation of data derived from this biobank. This is especially true for the gastrointestinal tissues that had lymphocytic and/or eosinophilic infiltrates. Previous tissue collection from both mares also had a similar infiltrate pattern in the gastrointestinal sections to the stallions described here (Burns et al., 2018). While this makes the direct comparison between male and female samples more consistent, it underscores the necessity to perform detailed phenotypic characterization of the tissues. The available phenotype information was expanded in this data set with the addition of ERG measurements. Further, this additional information enhances the ability to select tissue with the most normal function or, at minimum, document pathology that may affect future assays. In this case, the left retina from ECA_UCD_AH3 with an abnormal waveform was not used for the tissue bank, with the normal right retina retained.

The available whole-genome sequencing and SNP genotyping data from these horses to accompany the biobank and pending tissue-specific regulatory information will assist in future genomic investigations. In this sampling effort, the availability of a stallion closely related to the mare from which the reference genome is based (Twilight) is expected to provide high accuracy mapping and allow for more in depth investigations of achieved vs. predicted inbreeding. The intentional inbreeding that resulted in ECA_UCD_AH4 (by breeding record) was corroborated by the frequency and relatively large ROH as identified in the analysis of the SNP genotype data. This also highlights the limitation of this biobank, as it is restricted to a single breed. Future investigation may need to include more diverse biologic replicates to account for between breed variations in gene regulation (Li et al., 2012).

While the equine FAANG project has not progressed to the level of the ENCODE project in humans, important data have already been generated from the equine biobank

(ENCODE Project Consortium, 2012). Histone modifications as analyzed by chromatin immunoprecipitation sequencing have been established for eight tissues from the original two mares (Kingsley et al., 2019). The addition of repressive and enhancing epigenetic marks to the catalog of equine genome regulation is of great benefit to the equine research community. Sex-biased gene regulation is strongly intertwined with epigenetic regulation, especially as it relates to histone modifications (Tsai et al., 2009). This reinforces the need for the equine FAANG project to expand the existing biobank with male samples. This undertaking was no small feat and required the sacrifice of two additional animals. However, the benefit to future equine genomic studies is immense.

Although this study concludes the tissue collection arm of the FAANG project, the work to annotate the equine genome continues in earnest. This addition to the biobank expands the list of tissues available to the equine research community. Groups with specific interest in certain tissues are encouraged to contact the corresponding author for availability, as the “adopt-a-tissue” program successfully used with the original collection is also available for these male samples.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: <https://www.ebi.ac.uk/ena, PRJEB42407>.

ETHICS STATEMENT

The animal study was reviewed and approved by UC Davis Institutional Animal Care and Use Committee.

REFERENCES

- Baxter, G. M. (ed.). (2020). *Adams and Stashak's Lameness in Horses*. Hoboken, NJ: John Wiley & Sons.
- Burns, E. N., Bordbari, M. H., Mienaltowski, M. J., Affolter, V. K., Barro, M. V., Gianino, F., et al. (2018). Generation of an equine biobank to be used for Functional Annotation of Animal Genomes project. *Anim. Genet.* 49, 564–570. doi: 10.1111/age.12717
- ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature* 489, 57–74. doi: 10.1038/nature11247
- Fenner, K., Caspar, G., Hyde, M., Henshall, C., Dhand, N., Probyn-Rapsey, F., et al. (2019). It's all about the sex, or is it? Humans, horses and temperament. *PLoS ONE* 14:e0216699. doi: 10.1371/journal.pone.0216699
- Furr, M., and Reed, S. (eds.). (2015). *Equine Neurology*. Hoboken, NJ: John Wiley & Sons.
- Garbe, J. R., and Da, Y. (2008). *User Manual Version 2.4*. Minneapolis, MN: Department of Animal Science, University of Minnesota.
- Hida, N., Maeda, Y., Katagiri, K., Takasu, H., Harada, M., and Itoh, K. (2002). A simple culture protocol to detect peptide-specific cytotoxic T lymphocyte precursors in the circulation. *Cancer Immunol. Immunother.* 51, 219–228. doi: 10.1007/s00262-002-0273-7
- Kingsley, N. B., Kern, C., Creppe, C., Hales, E. N., Zhou, H., Kalbfleisch, T. S., et al. (2019). Functionally annotating regulatory elements in the equine genome using histone mark ChIP-Seq. *Genes* 11:3. doi: 10.3390/genes11010003
- Lee, S. K. (2018). Sex as an important biological variable in biomedical research. *BMB Rep.* 51, 167–173. doi: 10.5483/BMBRep.2018.51.4.034

AUTHOR CONTRIBUTIONS

Clinical examinations were performed by CF, SK, KK, and CD. Semen evaluation and cryopreservation performed by GD. Karyotyping performed by TR. Tissue collection performed by EH, AN, KK, NK, JT, EE, SP, AD, AF, JP, and CD. All histopathology reviewed by VA. Nuclei preparation performed by RB and FA. Cell culture and cryopreservation performed by MM. Genetic analysis performed by JP and RB. All authors contributed to manuscript preparation.

FUNDING

Funding was provided by the Grayson Jockey Club Foundation, USDA NRSP-8 and the UC Davis Center for Equine Health. Support for CF was provided by the National Institutes of Health (NIH) (L40 TR001136). None of the funding agencies had any role in the design of the study, analysis, interpretation of the data, or writing of the manuscript. Salary support for SP was provided by the Ann T. Bowling Fellowship at the UC Davis Veterinary Genetics Laboratory.

ACKNOWLEDGMENTS

The authors would like to thank all of the technicians at the Center for Equine Health and Anatomic Pathology.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2021.650305/full#supplementary-material>

- Li, M., Wu, H., Luo, Z., Xia, Y., Guan, J., Wang, T., et al. (2012). An atlas of DNA methylomes in porcine adipose and muscle tissues. *Nat. Commun.* 3, 1–11. doi: 10.1038/ncomms1854
- Marras, G., Gaspa, G., Sorbolini, S., Dimauro, C., Ajmone-Marsan, P., Valentini, A., et al. (2015). Analysis of runs of homozygosity and their relationship with inbreeding in five cattle breeds farmed in Italy. *Anim. Genet.* 46, 110–121. doi: 10.1111/age.12259
- Raimondi, E., Piras, F. M., Nergadze, S. G., Di Meo, G. P., Ruiz-Herrera, A., Ponsa, M., et al. (2011). Polymorphic organization of constitutive heterochromatin in *Equus asinus* (2n = 62) chromosome 1. *Hereditas* 148, 110–113. doi: 10.1111/j.1601-5223.2011.02218.x
- Raudsepp, T., Durkin, K., Lear, T. L., Das, P. J., Avila, F., Kachroo, P., et al. (2010). Molecular heterogeneity of XY sex reversal in horses. *Anim. Genet.* 41(Suppl. 2), 41–52. doi: 10.1111/j.1365-2052.2010.02101.x
- Raudsepp, T., Lee, E. J., Kata, S. R., Brinkmeyer, C., Mickelson, J. R., Skow, L. C., et al. (2004). Exceptional conservation of horse-human gene order on X chromosome revealed by high-resolution radiation hybrid mapping. *Proc. Natl. Acad. Sci. U.S.A.* 101, 2386–2391. doi: 10.1073/pnas.0308513100
- Sieck, R. L., Fuller, A. M., Bedwell, P. S., Ward, J. A., Sanders, S. K., Xiang, S. H., et al. (2020). Mandibulofacial dysostosis attributed to a recessive mutation of CYP26C1 in hereford cattle. *Genes* 11:1246. doi: 10.3390/genes11111246
- Tsai, H. W., Grant, P. A., and Rissman, E. F. (2009). Sex differences in histone modifications in the neonatal mouse brain. *Epigenetics* 4, 47–53. doi: 10.4161/epi.4.1.7288
- Will, T. R., Proaño, S. B., Thomas, A. M., Kunz, L. M., Thompson, K. C., Ginnari, L. A., et al. (2017). Problems and progress regarding sex bias

and omission in neuroscience research. *eNeuro* 4:ENEURO.0278-17.2017. doi: 10.1523/ENEURO.0278-17.2017

Zucker, I., and Beery, A. (2010). Males still dominate animal studies. *Nature* 465:690. doi: 10.1038/465690a

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Donnelly, Bellone, Hales, Nguyen, Katzman, Dujovne, Knickelbein, Avila, Kalbfleisch, Giulotto, Kingsley, Tanaka, Esdaile, Peng, Dahlgren, Fuller, Mienaltowski, Raudsepp, Affolter, Petersen and Finno. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Table S1: List of tissues collected for two adult Thoroughbred stallions. Asterisk (*) denotes tissues with nuclei isolated.

<u>Musculoskeletal</u> Cartilage <ul style="list-style-type: none"> • Stifle • Thoracic metacarpo-phalangeal • Pelvic metacarpo-phalangeal Coronary periople (thoracic & pelvic) Deep digital flexor tendon (thoracic & pelvic) Superficial digital flexor tendon (thoracic & pelvic) Suspensory ligament (thoracic & pelvic) Gluteal muscle Longissimus dorsi muscle* Sacrocaudalis lateralis muscle 3 rd Metacarpal/tarsal diaphysis Rib bone marrow Femur bone marrow Sesamoid <u>Integument</u> Coronary periople (thoracic & pelvic) Lamina (thoracic* & pelvic) Hoof wall (thoracic & pelvic) Neck skin Unpigmented skin Dorsum skin Gluteal adipose tissue* Abdominal adipose tissue Loin adipose tissue	<u>Cardiovascular</u> Left atrium free wall Left ventricle free wall* Right atrium free wall Right ventricle free wall Aortic valve Mitral valve Pulmonic valve Tricuspid valve <u>Respiratory</u> Trachea Larynx Cricothyroid dorsalis Left lung* Mediastinal/bronchiol lymph node* Epiglottis <u>Endocrine</u> Thyroid Adrenal cortex (left & right) Adrenal medulla (left & right) Pancreas <u>Urogenital</u> Kidney cortex (left* & right) Kidney medulla (left & right) Urinary bladder Testis (left & right) Right epididymis <ul style="list-style-type: none"> • Caput • Cauda • Corpus Bulbourethral gland Prostate Seminal vesicle Ampulla Ductus deferens Corpus spongiosum Corpus cavernosum	<u>Gastrointestinal</u> Liver (left lobe)* Spleen* Tongue Esophagus Stomach Duodenum Jejunum* Ileum Cecum Right dorsal colon* Right ventral colon Left dorsal colon Left ventral colon Small colon <u>Nervous System</u> Dur mater Parietal cortex* Occipital cortex Frontal cortex Temporal cortex Corpus callosum Thalamus Hypothalamus Pituitary Pons Cerebellar vermis Cerebellum lateral hemisphere* Spinal cord <ul style="list-style-type: none"> • C1 • C6 • T1* • T8 Sciatic nerve <u>Eye</u> Cornea Retina Iris
--	---	--

Table S2: Histopathology findings for two adult Thoroughbred stallions.

Acronyms: no significant findings (NSF), pelvic limb (PL), thoracic limb (TL), cervical (C), thoracic (T), lumbar (L)

Tissue	ECA_UCD_AH3	ECA_UCD_AH4
Skin Mid Neck	NSF	NSF
Skin Dorsum	NSF	NSF
Gluteal Adipose	NSF	NSF
Gluteal Muscle	NSF	NSF
Sacrocaudalis Muscle	NSF	NSF
Longissimus Dorsi Muscle	NSF	NSF
Right TL coronary band	NSF	NSF
Right TL & PL Lamina	NSF	NSF
Right PL suspensory lig.	NSF	NSF
Right PL Superficial Digital Flexor Tendon	NSF	NSF
Left TL & Right PL Deep Digital Flexor Tendon	NSF	NSF
Sciatic Nerve	NSF	
Liver	Increased bile ducts/oval cells; mild lymphocytic portal infiltrates;	Rare centrolobular and peripherolobular undefined round cell infiltrate and focal subcapsular fibrosis
Spleen	Active white pulp	NSF
Adrenal Cortex & Medulla	Markedly folded cortex; multifocal; vacuolar change of cortical cells	NSF
Kidney Cortex	NSF	NSF
Kidney Medulla	Mild lymphocytic infiltration	Mild lymphocytic infiltration
Larynx	NSF	Mild subepithelial lymphoplasmocytic infiltrate
Cricothyroid Dorsalis Muscle	Minimal interstitial lymphocytes	NSF
Pancreas	Sub capsular fibrosis	NSF
Thyroid	NSF	NSF
Mesenteric Lymph node	Marked cortical (active follicles) and paracortical hyperplasia	Cortical and paracortical hyperplasia
Tracheobronchial Lymph Node	Marked cortical (active follicles) and paracortical hyperplasia	Cortical and paracortical hyperplasia
Left Lung	Multifocal fibrosis with mild lymphoplasmacellular infiltrates, contracted bronchi with lymphoid infiltrates	Perivascular lymphoid infiltrate; contracted bronchioles
Heart Left Atrium	NSF	Focal perivascular lymphoid myocarditis
Heart Left Ventricle	NSF; 4cm	Small aggregates of undefined round cells; 4.2cm
Heart Right Atrium	NSF	NSF
Heart Right Ventricle	NSF; 4.2cm	Perivascular and interstitial lymph-plasmacytic infiltrate in coronary fat/epicardium; 5.2cm
Ventricular septum	NSF; 4.1cm	NSF; 4.5cm
Mitral Valve	NSF	NSF
Tricuspid Valve	NSF	NSF
Aortic Valve	NSF	NSF
Pulmonic Valve	NSF	NSF
Trachea	NSF	NSF
Tongue	Minimal multifocal perivascular lymphocytic glossitis	Minimal multifocal perivascular lymphocytic glossitis
Epiglottis	Minimal lymphoplasmacellular infiltrates in lamina propria	NSF
Esophagus	NSF	NSF

Stomach	Lymphocytic gastritis with a lymph follicle formation	NSF
Duodenum	Mild enteritis	Mild enteritis
Ileum	Mild enteritis	Mild enteritis
Cecum	Lympho-plasmacytic and eosinophilic inflammation with prominent lymphoid tissue and lymphocytic exocytosis	Lympho-plasmacytic and eosinophilic inflammation with prominent lymphoid tissue and lymphocytic exocytosis
Right Ventral Colon	Lympho-plasmacytic and eosinophilic inflammation with prominent lymphoid tissue and lymphocytic exocytosis	Lympho-plasmacytic and eosinophilic inflammation with prominent lymphoid tissue and lymphocytic exocytosis
Left Ventral Colon	Lympho-plasmacytic and eosinophilic inflammation with prominent lymphoid tissue and lymphocytic exocytosis	Lympho-plasmacytic and eosinophilic inflammation with prominent lymphoid tissue and lymphocytic exocytosis
Left Dorsal Colon	Lympho-plasmacytic and eosinophilic inflammation with prominent lymphoid tissue and lymphocytic exocytosis	Lympho-plasmacytic and eosinophilic inflammation with prominent lymphoid tissue and lymphocytic exocytosis
Right Dorsal Colon	Lympho-plasmacytic and eosinophilic inflammation with prominent lymphoid tissue and lymphocytic exocytosis	Lympho-plasmacytic and eosinophilic inflammation with prominent lymphoid tissue and lymphocytic exocytosis
Small Colon	NSF	NSF
Eye (retina, cornea, iris, sclera)	NSF; retina missing	NSF
Frontal Cortex	NSF	Solitary perivascular lymphoid cuff
Occipital Cortex	NSF	NSF
Parietal Cortex	NSF	NSF
Temporal cortex	NSF	Perivascular lymphoid cuff
Pituitary	NSF	NSF
Cerebellum Vermis	NSF	NSF
Cerebellum lateral hemisphere	NSF	NSF
Pons	NSF	Perivascular lymphoid cuff
Thalamus	NSF	NSF
Hypothalamus	NSF	Focal meningeal lymphoid infiltrate
Dura Mater	NSF	NSF
Corpus Callosum	NSF	NSF
Spinal Cord C1	NSF	NSF
Spinal Cord C6	NSF	NSF
Spinal Cord T8	NSF	NSF
Spinal Cord L1	NSF	NSF
Spinal Cord L6	NSF	NSF
Dorsal Root Ganglia C4/5	NSF	NSF
Dorsal Root Ganglia T9/10	NSF	NSF
Urinary Bladder	NSF	NSF
Urethra	NSF	NSF
Prostate	NSF	NSF
Right Testicle	NSF; spermatogenesis present	NSF; spermatogenesis present
Left Testicle	NSF; spermatogenesis present	NSF; spermatogenesis present

Right caput epididymis		NSF
Right cauda epididymis	lymphocytic perivascular epididymitis	NSF
Right corpus epididymis	NSF	Minimal interstitial lymphocytic infiltrates
Seminal vesicle	NSF	NSF
Bulbourethral gland	NSF	NSF
Ampulla	NSF	NSF
Ductus deferens	NSF	NSF
Urethral process/corpus spongiosum	NSF	Lymphocytic balanitis
Body of penis/corpus cavernosum	NSF	NSF
Metacarpal III	NSF	NSF
Sesamoid	NSF	NSF

Table S3. a- and b- wave amplitudes and implicit times for scotopic electroretinography waveforms of two Thoroughbred stallions

	Horse	ECA_UCD_AH3						ECA_UCD_AH4					
	Flash stimulus (cd.s/m2) - white	0.01		3		10		0.01		3		10	
	Eye	OD	OS	OD	OS	OD	OS	OD	OS	OD	OS	OD	OS
a-Wave	amplitude (uV)	16.7	9.0	29.2	91.4	17.2	151.0	5.8	4.5	144.0	110.1	189.3	136.0
	implicit time (ms)	12.8	16.4	12.8	11.8	6.7	10.8	22.5	14.5	14.1	14.6	13.8	14.6
b-Wave	amplitude (uV)	230.8	186.8	358.1	200.8	614.0	375.8	245.6	305.7	364.8	498.1	411.4	444.7
	implicit time (ms)	83.5	91.6	38.9	35.3	69.6	65.5	76.8	67.8	57.7	58.5	61.4	59.2

Table S4: Epidydimal sperm post-thaw analysis for two adult Thoroughbred stallions

	ECA_UCD_AH3	ECA_UCD_AH4
Concentration (million/ml)	323.3	455.5
Number of straws		
Rapid progressive%	39.11	52.75
Slow progressive%	0.0	0.0
Non-progressive%	9.43	13.34
Immotile%	51.46	33.91
VCL (µm/sec)	110.2	100.5
VAP (µm/sec)	56.75	51.16
VSL(µm/sec)	43.03	36.59
LIN	0.39	0.36
STR	0.75	0.71
BCF (hertz)	26.23	22.58
ALH (µm)	4.28	4.49

	Table S5: Synovial fluid analysis for two adult Thoroughbred stallions					
	ECA_UCD_AH3			ECA_UCD_AH4		
	Left tarso-crural	Right middle carpal	Right radio carpal	Left tarso-crural	Left middle carpal	Left radio carpal
Gross appearance	Yellow/slight haze	Yellow/clear	Yellow/hazy	Yellow/clear	Yellow/hazy	Yellow/clear
Total protein (g/dL)	1.2	1.5	1.5	1.2	2.1	1.7
Mucin clot	Very good	Very good	Very good	Very good	Very good	Very good
Total RBC's (/μL)	<30,000	<30,000	<30,000	Rare	Rare	Rare
Total nucleated cells (/μL)	180	150	180	NA	NA	NA
Neutrophil %	0	1	2	0	1	0
Small mononuclear%	85	77	63	72	78	60
Large mononuclear%	15	22	35	28	21	40
Eosinophil%	0	0	0	0	0	0
Interpretation	No cytological abnormalities	No cytological abnormalities	No cytological abnormalities	No cytological abnormalities	No cytological abnormalities	No cytological abnormalities

	Table S6: Cerebrospinal fluid analysis for two adult Thoroughbred stallions	
	ECA_UCD_AH3	ECA_UCD_AH4
Gross appearance	Clear	Clear
Total protein (mg/dL)	72	88
Total RBC's (/μL)	6	<1
Total nucleated cells (/μL)	3	1
Neutrophil %	0	0
Small mononuclear%	95	85
Large mononuclear%	5	15
Eosinophil%	0	0
Interpretation	No cytological abnormalities	No cytological abnormalities

Table S7. Runs of homozygosity (ROH) and inbreeding estimate (F_{ROH}) for combined mare and stallion FAANG genome data sets				
Length of ROH (Mb)	ECA_UCD_AH1	ECA_UCD_AH2	ECA_UCD_AH3	ECA_UCD_AH4
0 to 4	466	462	453	220
4 to 8	24	26	38	29
8 to 16	7	11	6	16

16 to 32	3	4	1	11
> 32	0	0	0	4
ROH Genome	765915572	838359933	758961216	1009849409
F _{ROH}	0.336	0.368	0.333	0.444

Table S8. Microsatellite genotyping of cryopreserved fibroblast cell line DNA from two adult Thoroughbred stallions compared to whole blood derived genomic DNA

	Type			
	Whole Blood	Fibroblast	Whole Blood	Fibroblast
Locus	ECA_UCD_AH3	ECA_UCD_AH3	ECA_UCD_AH4	ECA_UCD_AH4
<i>AHT4</i>	JK	JK	K	K
<i>AME</i>	YX	YX	YX	YX
<i>ASB2</i>	N	N	O	O
<i>HMS3</i>	I	I	MP	MP
<i>HMS7</i>	M	M	LM	LM
<i>HTG4</i>	K	K	KM	KM
<i>LEX33</i>	MQ	MQ	Q	Q
<i>HMS2</i>	L	L	L	L
<i>AHT5</i>	J	J	KN	KN
<i>ASB17</i>	GM	GM	NR	NR
<i>ASB23</i>	J	J	J	J
<i>HMS6</i>	MP	MP	MP	MP
<i>HTG10</i>	IO	IO	R	R
<i>LEX3</i>	H	H	N	N
<i>VHL20</i>	I	I	IL	IL

Table S9. Color genotyping for two adult Thoroughbred stallions

	ECA_UCD_AH3	ECA_UCD_AH4
Red Factor	E/e	E/e
Agouti	A/a	A/a
Cream	N/N	N/N
Pearl	N/N	N/N
Silver	N/N	N/N
Dun	nd2/nd2	nd2/nd2
Champagne	N/N	N/N
Lethal White Overo	N/N	N/N
Sabino 1	N/N	N/N
Dominant White (W5, W10, W20, W22)	N/W20	W20/W20
Splashed White (SW1, SW3, SW5, SW6)	N/N	N/N
Splashed White (SW2, SW4)	N/N	N/N
Tobiano	N/N	N/N
Leopard	N/N	N/N
Patern-1	N/N	N/N
Gray	N/N	N/G